

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:18:28 ; Search time 64.32 Seconds

(without alignments)  
11.696 Million cell updates/sec

Title: US-09-331-631a-39

Perfect score: 54  
Sequence: 1 CXXCXXXXXXXXXXCXXC 22

Scoring table: BLOSUM62DX  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_36.\*  
1: /SIDSL/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	27	20	Y36499
2	54	100.0	31	21	Y70731
3	54	100.0	36	4	P30262
4	54	100.0	36	4	P30263
5	54	100.0	39	17	W05340
6	54	100.0	39	17	W05341
7	54	100.0	39	17	W05342
8	54	100.0	39	20	Y07909
9	54	100.0	44	17	R98208
10	54	100.0	44	21	Y64770
11	54	100.0	57	21	Y57813
12	54	100.0	59	21	Y57812

13	54	100.0	62	21	Y57810	Human metallothion
14	54	100.0	63	21	Y57815	Sea urchin metallo
15	54	100.0	66	21	Y64780	Human 5' EST relat
16	54	100.0	70	21	Y75953	Murine skin cell p
17	54	100.0	72	19	W50389	Snake venom platelet
18	54	100.0	72	19	W50452	Snake venom platelet
19	54	100.0	72	19	W46214	Snake venom platelet
20	54	100.0	73	20	Y35935	Snake venom platelet
21	54	100.0	75	18	W21583	Extended human sec
22	54	100.0	76	20	Y02761	Alzheimer's diseas
23	54	100.0	76	21	Y68907	Human secreted pro
24	54	100.0	80	17	W05343	A mouse MDNM-2 pro
25	54	100.0	86	20	Y36498	Calisoga spider ve
26	54	100.0	92	21	Y69209	Fragment of human
27	54	100.0	93	20	Y36164	Amino acid sequenc
28	54	100.0	93	20	Y36211	Human secreted pro
29	54	100.0	100	21	Y65659	Human secreted pro
30	54	100.0	104	21	Y65661	C. elegans insulin
31	54	100.0	105	21	Y65660	C. elegans insulin
32	54	100.0	106	21	Y65655	C. elegans insulin
33	54	100.0	112	21	Y65658	C. elegans insulin
34	54	100.0	118	21	Y44985	Human epidermal pr
35	54	100.0	118	21	Y65662	C. elegans insulin
36	54	100.0	124	19	W56732	Nucleolus specific
37	54	100.0	125	12	R13329	He4 epidiolymis-spe
38	54	100.0	125	19	W81779	Human HE4 protein.
39	54	100.0	128	21	Y44987	Human epidermal pr
40	54	100.0	149	8	P70057	Human insulin rece
41	54	100.0	150	8	P70058	Human epidermal gr
42	54	100.0	165	12	R10533	Prod. of pwc4812 u
43	54	100.0	169	20	Y60558	Human normal blad
44	54	100.0	170	20	Y29215	Amino acid sequenc
45	54	100.0	233	21	Y74791	Neisseria meningit

#### ALIGNMENTS

RESULT 1	ID	Y36499	standard; Protein: 27 AA.
XX	XX	Y36499;	
XX	AC		
XX	AC		
XX	DT	17-SEP-1999	(first entry)
XX	DE		Fragment of human secreted protein encoded by gene 27.
XX	KW		Human; secreted protein; cancer; tumour; developmental abnormality;
XX	KW		foetal deficiency; blood disorder; immune system disorder; inflammation;
XX	KW		autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
XX	KW		schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
XX	KW		atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
XX	KW		digestive disorder; endocrine disorder; infection; AIDS.
OS	XX		Homo sapiens.
XX	PN	W09931117-A1.	
XX	PD	24-JUN-1999.	
XX	PF	17-DEC-1998;	98WO-US27059.
XX	PR	19-DEC-1997;	97US-0068369.
XX	PR	18-DEC-1997;	97US-0068006.
XX	PR	18-DEC-1997;	97US-0068007.
XX	PR	18-DEC-1997;	97US-0068008.
XX	PR	18-DEC-1997;	97US-0068053.
XX	PR	18-DEC-1997;	97US-0068054.
XX	PR	18-DEC-1997;	97US-0068057.
XX	PR	18-DEC-1997;	97US-0068064.
XX	PR	18-DEC-1997;	97US-0070923.
XX	PR	19-DEC-1997;	97US-0068169.

PR 19-DEC-1997; 97US-0068365.  
 PR 19-DEC-1997; 97US-0068367.  
 PR 19-DEC-1997; 97US-0068368.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;  
 PI Florence K, Greene JM, Janet F, Kyaw H, Moore PA;  
 PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;  
 PI Yu G;  
 XX  
 DR WPI; 1999-418749/35.  
 XX  
 PT New isolated human genes encoding secreted polypeptides  
 PS  
 XX Disclosure; Page 466; 537pp; English.  
 PS  
 XX X97916 to X98029 represent 110 isolated human secreted protein genes.  
 CC Y36224 to Y36727 represent the secreted proteins encoded by the 110  
 CC human genes. The genes and their corresponding secreted polypeptides are  
 CC useful for preventing, treating or ameliorating medical conditions,  
 CC e.g. by protein or gene therapy. Also pathological conditions can be  
 CC diagnosed by determining the amount of the new polypeptides in a sample  
 CC or by determining the presence of mutations in the new genes. Specific  
 CC uses are described for each of the 110 genes, based on which tissues they  
 CC are most highly expressed in, and include developing products for the  
 CC diagnosis or treatment of cancer, tumours, developmental abnormalities  
 CC and foetal deficiencies, blood disorders, diseases of the immune system,  
 CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive  
 CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin  
 CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney  
 CC disorders, digestive/endocrine disorders, infections and AIDS. The  
 CC polypeptides are also useful for identifying their binding partners.  
 CC The sequences given in X97907 to X97915 and Y36223 are used in the  
 CC exemplification of the present invention.  
 XX  
 SQ Sequence 27 AA:  
 Query Match 100.0%; Score 54; DB 20; Length 27;  
 Best Local Similarity 18.2%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 18; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXXXXXXXXX 22  
 I:::|:::|:::|:::|:::|  
 Db 4 cpccclprgscgrqrcfsc 25  
 RESULT 2  
 Y70731  
 ID Y70731 standard; protein: 31 AA.  
 XX  
 AC Y70731:  
 XX  
 DT 24-JUL-2000 (first entry)  
 XX  
 DE Wnt antagonist protein consensus sequence-1.  
 XX  
 DE Wnt antagonist; contraceptive; contraceptive vaccine; oocyte development;  
 KW female primate contraception; oocyte viability.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 2  
 FT /label= Unknown  
 FT /note= "Xaa may be 9 amino acids in length; some  
 FT amino acids may be absent."  
 FT Misc-difference 4  
 FT /label= Unknown  
 FT /note= "Xaa may be 42 amino acids in length; some  
 FT amino acids may be absent"  
 FT Misc-difference 14

FT /label= Unknown  
 FT Misc-difference 15  
 FT /label= Unknown  
 FT Misc-difference 16  
 FT /label= Unknown  
 FT Misc-difference 17  
 FT /label= Unknown  
 FT Misc-difference 18  
 FT /label= Unknown  
 FT Misc-difference 19  
 FT /label= Unknown  
 FT Misc-difference 21  
 FT /label= Unknown  
 FT /note= "Xaa may be 10 amino acids in length; some  
 FT amino acids may be absent"  
 FT Misc-difference 23  
 FT /label= Unknown  
 FT Misc-difference 24  
 FT /label= Unknown  
 FT Misc-difference 25  
 FT /label= Unknown  
 FT Misc-difference 27  
 FT /label= Unknown  
 FT /note= "Xaa may be 7 amino acids in length; some  
 FT amino acids may be absent"  
 FT Misc-difference 29  
 FT /label= Unknown  
 FT /note= "Xaa may be 27 amino acids in length; some  
 FT amino acids may be absent"  
 FT Misc-difference 31  
 FT /label= Unknown  
 FT /note= "Xaa may be 13 amino acids in length; some  
 FT amino acids may be absent"  
 XX  
 PN WO200021555-A1.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 13-OCT-1999; 99WO-US23640.  
 XX  
 PR 15-OCT-1998; 98US-0104355.  
 XX  
 PA (HARD) HARVARD COLLEGE.  
 XX  
 PI McMahon AP, Parr BA, Vaino S;  
 PI  
 DR WPI; 2000-317845/27.  
 XX  
 PT Contraceptive composition for inhibiting oocyte development in a female  
 PT primate comprises a Wnt polypeptide antagonist  
 XX  
 PS Claim 12; Page 44; 57pp; English.  
 XX  
 CC The patent discloses a method of female primate contraception comprising  
 CC administering an antagonist of a Wnt polypeptide, inhibiting oocyte  
 CC development. Wnt polypeptides are useful for promotive maturation of an  
 CC immature oocyte. Wnt polypeptides are also useful for increasing the  
 CC number of mature oocytes and to enhance oocyte viability. The present  
 CC peptide is a consensus sequence of Wnt antagonist which inhibits the  
 CC physiological activity of a Wnt polypeptide. Antagonistic polypeptides  
 CC may contain a cysteine-rich domain.  
 XX  
 SQ Sequence 31 AA:  
 Query Match 100.0%; Score 54; DB 21; Length 31;  
 Best Local Similarity 63.6%; Pred. No. 1.1e+02;  
 Matches 14; Conservative 8; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXXXXXXXXX 22  
 I:::|:::|:::|:::|:::|  
 Db 5 ccccccccccccccccc 26

```

RESULT 3
ID P30262 standard; peptide: 36 AA.
XX
AC P30262;
XX
DT 25-APR-1992 (first entry)
XX
DE Sequence of peptide used to vaccinate against E. coli enterotoxin(s).
XX
KW Vaccine; enterotoxin; diarrhoea; immunogen.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..18
PN /label= Peptide P
PD EP93652-A.
XX
XX 09-NOV-1983.
XX
PE 26-APR-1983; 83EP-0072336.
XX
PR 26-APR-1982; 82FR-0007179.
XX
PA (INSP ) INST PASTEUR.
XX (CNRS ) CENT NAT RECH SCT.
XX
PI Tartar A, Duflot E, Boquet P;
XX
DR WPI; 1983-816301/46.
XX
PT Peptide(s) used to vaccinate against E. coli enterotoxin(s) -
PT contg. e.g. asparagine threonine phenylalanine tyrosine cysteine
PT cysteine glutamic acid leucine cysteine asparagine
XX
PS Claim 1; Page 40; 50pp; French.
XX
CC The inventors claim peptides of formula (P)n (see FT; see also
CC P30263) having 4n-18n amino acids and pref. being laevorotatory
CC (where n is 1 or 2). In P30262 and P30263, N=2. When n is 2, the
CC peptide comprises two peptide sequences P1 which may be the same or
CC different, each having 4-18 amino acids chosen from the peptide P SQ
CC in P30262 or P30263. The two P sequences may be joined (a) by a
CC disulphide bond or (b) by a bond formed between a carboxyl gp. of
CC one sequence of an amino gp. of the other.
XX
SQ Sequence 36 AA;

Query Match 100.0%; Score 54; DB 4; Length 36;
Best Local Similarity 18.2%; Pred. No. 1.3e+02;
Matches 4; Conservative 18; Mismatches 0; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXXXCXXC 22
|::|::|::|::|::|::|
Db 6 celcnpacagcnytfyceic 27

RESULT 4
ID P30263 standard; peptide: 36 AA.
XX
AC P30263;
XX
DT 25-APR-1992 (first entry)
XX
DE Sequence of peptide used to vaccinate against E. coli enterotoxin(s).
XX
KW Vaccine; enterotoxin; diarrhoea; immunogen.

```

```

XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Misc-difference 1..18
PN /label= Peptide P
PD EP93652-A.
XX
XX 09-NOV-1983.
XX
PE 26-APR-1983; 83EP-0072336.
XX
PR 26-APR-1982; 82FR-0007179.
XX
PA (INSP ) INST PASTEUR.
XX (CNRS ) CENT NAT RECH SCT.
XX
PI Tartar A, Duflot E, Boquet P;
XX
DR WPI; 1983-816301/46.
XX
PT Peptide(s) used to vaccinate against E. coli enterotoxin(s) -
PT contg. e.g. asparagine threonine phenylalanine tyrosine cysteine
PT cysteine glutamic acid leucine cysteine asparagine
XX
PS Claim 1; Page 40; 50pp; French.
XX
CC The inventors claim peptides of formula (P)n (see FT; see also
CC P30263) having 4n-18n amino acids and pref. being laevorotatory
CC (where n is 1 or 2). In P30262 and P30263, N=2. When n is 2, the
CC peptide comprises two peptide sequences P1 which may be the same or
CC different, each having 4-18 amino acids chosen from the peptide P SQ
CC in P30262 or P30263. The two P sequences may be joined (a) by a
CC disulphide bond or (b) by a bond formed between a carboxyl gp. of
CC one sequence of an amino gp. of the other.
XX
SQ Sequence 36 AA;

Query Match 100.0%; Score 54; DB 4; Length 36;
Best Local Similarity 18.2%; Pred. No. 1.3e+02;
Matches 4; Conservative 18; Mismatches 0; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXXXCXXC 22
|::|::|::|::|::|::|
Db 6 celcypacagcnytfyceic 27

RESULT 5
ID W05340 standard; peptide: 39 AA.
XX
AC W05340;
XX
DT 15-APR-1997 (first entry)
XX
DE Callisoga spider venom peptide A, used as insecticide.
XX
KW Callisoga: spider; venom; insecticide; recombinant; baculovirus; pest;
KW tobacco budworm; Heliothis virescens; low mammalian toxicity.
XX
OS Callisoga sp.
XX
PN W09625041-A1.
XX
PD 22-AUG-1996.
XX
PF 16-FEB-1996; 96WO-US02030.
XX
PR 17-FEB-1995; 95US-0390882.
XX

```



RESULT	8	
Y07909		
ID	Y07909	standard; Protein; 39 AA.
AC	Y07909;	
XX		
DT	06-JUL-1999	(first entry)
XX		
DE		Human secreted protein fragment encoded from gene 58.
XX		
KW		Human; secreted protein; treatment; prevention; protein therapy; AIDS;
KW		gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;
KW		developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW		immune system disease; autoimmune disease; hepatic disease; lymphoma;
KW		renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia
KW		pulmonary disorder; prostate disease; skeletal; cardiac; muscle disorder
KW		pulmonary disorder; transplant rejection; osteoclast; osteoporosis;
XX		arthritis; malignancy; digestive; endocrine; infection.
OS		Homo sapiens.
XX		
PN	W09918208-A1.	
XX		
PD	15-APR-1999.	
XX		
PE	01-OCT-1998;	98MO-US20775.
XX		
PR	02-OCT-1997;	97US-0060884.
PR	02-OCT-1997;	97US-0060833.
PR	02-OCT-1997;	97US-0060836.
PR	02-OCT-1997;	97US-0060837.
PR	02-OCT-1997;	97US-0060838.
PR	02-OCT-1997;	97US-0060839.
PR	02-OCT-1997;	97US-0060843.
PR	02-OCT-1997;	97US-0060862.
PR	02-OCT-1997;	97US-0060866.
PR	02-OCT-1997;	97US-0060874.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
PI	Cartier KC,	Duan DR, Endress GA, Feng P, Ferrie AM,
PI	Florence KA,	Greene JM, Janat F, Lafleur DW, Ni J;
PI	Rosen CA,	Ruben SM, Shi Y, Young P, Yu G;
DR	WPI: 1999-264022/22.	
DR	N-PSDB: X37508.	
XX		
PT		New isolated human genes and the secreted polypeptides they encode
XX		
PS	Claim 1b; Page 306; 368pp;	English.
XX		
CC	This invention describes novel isolated human genes and the secreted	
CC	proteins they encode. The products of the invention are useful for	
CC	preventing, treating or ameliorating medical conditions, e.g. by protein	
CC	or gene therapy. Also pathological conditions can be diagnosed by	
CC	determining the amount of the new polypeptides in a sample or by	
CC	determining the presence of mutations in the new polynucleotides.	
CC	Specific uses are described for each of the 101 polynucleotides, based on	
CC	which tissues they are most highly expressed in, and include developing	
CC	products for the diagnosis or treatment of cancer, tumours,	
CC	neurodegenerative disorders, developmental abnormalities and fetal	
CC	deficiencies, blood disorders, leukemias, diseases of the immune system,	
CC	autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,	
CC	allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate	
CC	disease, skeletal or cardiac muscle disorders, pulmonary disorders,	
CC	transplant rejection, disorders involving osteoclasts such as	
CC	osteoporosis, arthritis or malignancies, digestive/endocrine disorders,	
CC	infections and AIDS. The human secreted proteins of the invention are	
CC	represented in X37451-X37552.	
XX		
Sequence	39 AA;	
10		

```

Query Match          100.0%; Score 54; DB 20; Length 39;
Best Local Similarity 18.2%; Pred. No. 1.4e+02;
Matches 4; Conservative 18; Mismatches 0; Indels 0; Gaps 0.

QY 1 CXXCXXXXXXXXXXXXCXXC 22
    ||:::||||:::||||:::|:::|
Db 6 cfraccmcslsgllnltcsc 27

```

```

RESULT      9
R98208
ID          R98208 standard; Protein; 44 AA.
XX
AC          R98208;
XX
DT          30-DEC-1996 (first entry)
XX
DE          Nucleotide used in production of MSH/Momulv chimeric sequence.
XX
KW          Moloney murine leukaemia virus; gp70; 4070A retrovirus; retrovirus;
KW          10A1 murine leukaemia virus; NZB-9-1 murine leukaemia virus;
KW          polytropic MK27 provirus; targeted drug delivery; gene therapy;
KW          single chain antibody; envelope protein; ss.
XX
OS          Synthetic.
XX
PN          WO9630504-A1.
PD          03-OCT-1996.
XX
PF          22-MAR-1996; 96WO-US03908.
XX
PR          24-MAR-1995; 95US-0409648.
XX
PA          (GENE-) GENETIC THERAPY INC.
PA          (OYSC-) UNIV SOUTHERN CALIFORNIA.
XX
PI          Anderson W, Chiang YL, Januszewski M, Mackrell AJ;
PI          Zhao Y;
XX
DR          WPI; 1996-455352/45.
XX
PT          Cell-targeted retroviral vector particles - having envelope protein
PT          modified with targetting polypeptide
XX
PS          Example 2; Page 36; 73pp; English.
XX
CC          Cell targeted retroviral vector particles can be used in gene
CC          therapy to deliver a heterologous gene to a target cell for
CC          expression of a heterologous polypeptide in that cell. The cell
CC          targeted retroviral vector particles comprise an envelope protein
CC          which is modified to contain a targetting polypeptide (a single chain
CC          antibody), or in the case of moloney murine leukaemia virus
CC          (Momulv), alpha melanotropin-stimulating hormone (MSH). Two
CC          oligonucleotides (R98207, R98208) were used to substitute sequences in
CC          Momulv for MSH sequences. This oligonucleotide was used to replace
CC          residues G80-P88 of Momulv envelope protein (see W04248).
XX
SQ          Sequence      44 AA;

Query Match      100.0%; Score 54; DB 17; Length 44;
Best Local Similarity 18.2%; Pred. NO. 1.5e+02;
Matches      4; Conservative 18; Mismatches      0; Indels      0; Gaps      0;

QY      1 CXXXCXXXXXXXXXXXXXCC 22
        |::|::|::|::|::|::|
Db      15 caagcgcgtattaacctcctc 36

RESULT      10

```

XX	Y64770 standard; Protein; 44 AA.
XX	
AC	Y64770;
XX	
DT	01-FEB-2000 (first entry)
XX	
DE	Human 5' EST related polypeptide seq ID NO:931.
XX	
KW	Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW	gene therapy; chromosome mapping; upstream regulatory sequence;
KW	forensic; location; development; protein synthesis; stability;
KW	regulation; identification.
XX	
OS	Homo sapiens.
XX	
PN	W09953051-A2.
XX	
PD	21-OCT-1999.
XX	
PF	09-APR-1999; 99WO-IB00712.
XX	
PR	09-APR-1998; 98US-0057719.
PR	28-APR-1998; 98US-0069047.
PA	(GEST ) GENSET.
PI	Dumas Milne Edwards J, Duclert A, Giordano J;
DR	WPI; 2000-038446/03.
DR	N-PSDB; 242384.
PT	Novel secreted protein 5' expressed sequence tag sequences used in
XX	diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX	
PS	Claim 3; Page 637; 837pp; English.
XX	
CC	242265 to 243075 represent novel 5' expressed sequence tag (EST)
CC	representing the EST-related proteins corresponding to Y6451 to Y65438
CC	represent the EST-related proteins corresponding to 242265 to 243052.
CC	The 5' ESTs can be used for producing secreted human gene products.
CC	They can be used to identify and isolate 5' untranslated regions (UTRs)
CC	and upstream regulatory regions which control the location, development
CC	stage, rate, and quantity of protein synthesis, as well as stability of
CC	mRNA. The ESTs are also useful as probes for chromosome mapping, and to
CC	obtain full length cDNA clones. The ESTs can also be used in forensic
CC	procedures to identify individuals, or in diagnostic procedures to
CC	identify individuals having genetic diseases resulting from abnormal
CC	gene expression. The products may also be used in gene therapy protocols.
CC	The nucleic acids encoding signal peptides can be used for directing
CC	extracellular secretion of a polypeptide or the insertion of a
CC	polypeptide into a membrane, or importing a polypeptide into a cell.
CC	The proteins encoded by the EST sequences may be useful in treating a
CC	variety of human conditions. Secreted proteins have therapeutic value,
CC	and the identification of new secreted proteins is valuable. 242249 to
CC	242264 and Y64644 to Y64650 represent sequences used in the
CC	exemplification of the present invention.
XX	
XX	Sequence 44 AA:

```

Query Match      100.0%; Score 54; DB 21; Length 44;
Best Local Similarity 18.2%; Pred. No. 1.5e+02;
Matches 4; Conservative 18; Mismatches 0; Indels 0; Gaps 0;

QY      1 CXXXCXXXXXXXXXXXXCXXC 22
        |:::|:::|:::|:::|
Db       15 CYSVCVSVPYGMYLVCVCVC 36

RESULT 11
Y57813
ID      Y57813 standard; protein; 57 AA.
XX

```

AC Y57813;  
XX  
DT 22-MAR-2000 (first entry)  
XX  
DE Crab metallothionein Class I amino acid sequence.  
XX  
KW Metallothionein; metal recovery; remediation; heavy metal;  
KM precious metal; phycochelatin; green algae; Chlamydomonas reinhardtii.  
XX  
OS Eubrachyura.  
XX  
PN WO9960838-A1.  
XX  
PD 02-DEC-1999.  
XX  
PF 28-MAY-1999; 99WO-USJ2007.  
XX  
PR 28-MAY-1998; 98US-0087374.  
XX  
PA (OHIS ) UNIV OHIO STATE RES FOUND.  
XX  
PI Sayre RT, Trajna SJ;  
XX  
DR WPI; 2000-086646/07.  
XX  
PT Novel method for metal recovery, remediation and separation -  
XX  
PS Disclosure; Page 6; 86pp; English.  
XX  
CC The present invention describes a transgenic algal cell (I) of the  
CC genus Chlamydomonas comprising reproductive genetic material comprising  
CC a nucleotide sequence capable of expressing chicken type I  
CC Metallothionein. Also described is a method of removing metal from  
CC an aqueous medium containing at least one dissolved or suspended  
CC metal. The transgenic algae are used for the selective separation of  
CC metals, particularly the separation of precious and desirable metals  
CC such as gold and uranium, from other metals such as cadmium, zinc and  
CC copper. The method can be used to facilitate the selective recovery of  
CC precious and rare metals from mineral sources where aqueous media can  
CC be used, such as in natural surface water flows, ground water and where  
CC water may be introduced. The method is suitable for well-drilling,  
CC soil and water remediation arts, mining fields, and industrial  
CC engineering. The present sequence represents a Class I metallothionein  
CC given in the present invention.  
XX  
SQ Sequence 57 AA;

Query Match	100.0%	Score 54:	DB 21:	Length 57:
Best Local Similarity	18.2%	Pred. NO. 1.9e+02:		
Matches	4:	Conservative	18:	Mismatches 0:
				Indels 0:
				Gaps 0:
Qy	1	CXXXCXXXXXXXXXXCXXC	22	
		::: ::: ::: ::: ::: :::		
Db	16	ctgctcctscrcpccgcsgc	37	
RESULT	12			
ID	y57812			
	y57812	standard; protein; 59 AA.		
XX				
AC	y57812;			
XX				
DT	22-MAR-2000	(first entry)		
XX				
DE	Trout metallothionein Class I amino acid sequence.			
XX				
KW	Metallothionein; metal recovery; remediation; heavy metal;			
XX	precious metal; phycochelatin; green algae; Chlamydomonas reinhardtii.			
OS	Salmo sp.			
XX				
PN	W09960838-A1			



